IDENTIFYING THE COGNITIVE DECREMENTS CAUSED BY HIV

Diane Damos
Richard S. John
Elizabeth S. Parker
Alexandra M. Levine

Interim Technical Report
June, 1994

University of Southern California

IDENTIFYING THE COGNITIVE DECREMENTS CAUSED BY HIV

Diane Damos

Richard S. John

Elizabeth S. Parker

Alexandra M. Levine

Interim Technical Report

June, 1994

University of Southern California

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	10 June 1994	Interim Aug. 1990- June 1994

4. TITLE AND SUBTITLE

Identifying the Cognitive Decrements Caused by HIV

5. FUNDING NUMBERS

N00014-90-J-4079

6. AUTHOR(S)

Diane L. Damos
Richard S. John
Elizabeth S. Parker
Alexandra M. Levine

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute of Safety and Systems Management
University of Southern California
Los Angeles, CA, 90045-0021

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This study had two purposes. The first was to determine the disease stage at which cognitive decrements caused by HIV become detectable. The second was to compare the sensitivity of information processing tests to neuropsychological instruments for detecting cognitive deficits caused by HIV. The study design initially had five groups: an asymptomatic group (Walter Reed Stages 1,2, and 3), a symptomatic group (Walter Reed Stages 4 and 5), a homosexual, HIV-negative control group, a heterosexual, HIV-negative control group, and a group receiving anti-retroviral medication.

After the data had been collected, the method of administering the Merieux, one of the staging criteria, was found to be unreliable. Consequently, the subjects were reclassified and the analyses focused on two groups: the asymptomatic group, which was now defined as subjects in either Walter Reed Stage 1 or 2, and the homosexual control group. Few differences were found between the two groups. The primary differences appeared to involve verbal fluency and verbal episodic memory.

14. SUBJECT TERMS

HIV
Cognitive deficits
Information processing
Neuropsychological assessment

15. NUMBER OF PAGES

76

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Many individuals contributed to the success of this project beside the authors. The authors take this opportunity to recognize them:

Dr. Deborah Steese-Seda made the initial arrangements for the subjects and helped organize the testing process;

Dr. Steven Whipple conducted much of the neuropsychological testing and supervised the day-to-day running of the laboratory;

Dr. Daniel Salzer spent many hours contacting various awareness groups and clinics to recruit subjects. He also performed the neuropsychological testing;

Mr. Olukoyode Olufinboba programmed the data acquisition software and performed much of the data analysis;

Mr. Alok Sharma also helped with the data analysis;

Mr. Nelson DeGuzman performed the physical examinations;

Mr. Brian Buller and Ms. Denyse Worley conducted the information processing tests;

Ms. Gina Galante and Ms. Jennifer Nikcevic helped edit this manuscript.

IDENTIFYING COGNITIVE DECREMENTS CAUSED BY HIV

INTRODUCTION

Cognitive deterioration is one of the hallmarks of HIV infection. Today, there is little, if any, debate that HIV infection has adverse effects on cognitive processes at the advanced stages of the disease (McAllister et al., 1992; Poutiainen et al., 1993). Instead, the controversy centers on HIV+ asymptomatic individuals. Because involvement of the central nervous system by HIV appears to begin at the time of initial infection (Bornstein et al., 1991; Lunn et al., 1991; McAllister et al., 1992), cognitive decline theoretically could begin during the asymptomatic stage.

The literature addressing cognitive decline in asymptomatic individuals is decidedly mixed. Several seminal articles published in 1987 and 1988 appeared to demonstrate cognitive decline in a large proportion of asymptomatic individuals (Grant et al., 1987; Saykin, et al., 1988). The majority of studies published immediately subsequent to these reports failed to replicate the initial results (e.g., McArthur et al., 1989), and the original studies were criticized on a number of methodological grounds. Nevertheless, more recent studies again have reported a decline in cognitive function in asymptomatic individuals (Bornstein et al., 1992; Lunn et al., 1991; Sinforiani et al., 1991).

The conflicting results concerning the decline in cognitive

processes in asymptomatic individuals may be attributed partially to variations in methodology and dependent measures. Specifically, recent studies have placed greater emphasis on the use of reaction time, which is measured to millisecond accuracy, than on more traditional dependent measures, such as percent correct or time to completion, which is measured in seconds. Because various methodological issues and the use of reaction time measures strongly influenced the design and procedures of the current study, the following two sections review a number of methodological issues and the use of reaction time as a dependent measure.

Methodological Issues

Even a casual reading of the earlier literature studying the effects of HIV on cognitive processes reveals a surprising number of methodological problems. For the majority of these problems--such as the recruitment of non-native English speakers--differential enrollment could adversely affect the performance of the group with the highest proportion of problematic subjects. We describe 12 methodological problems that we feel are the most important for HIV studies. These problems are grouped into three categories: screening, sampling, and methodological issues specific to HIV research.

Screening. One of the most difficult aspects of subject screening involves alcohol and substance abuse. Effective initial screening relies on accurate self report, a trait that is questionable in heavy substance abusers. Toxicological screens may

be used in addition to self report, but these are expensive, particularly serum screening, which requires specialized personnel to draw and handle the blood. Urine screening, which is less costly, only indicates if the subject has had a particular substance in the last few days or weeks, but it does not provide quantitative information about the actual level of psychoactive drugs. Serum screening is better for that purpose. Regardless of the type of screening administered, any delay between the toxicological screening and the cognitive assessment may allow the subject to abuse substances that adversely affect cognitive performance while appearing toxicologically negative.

Enrolling subjects with a prior history of psychiatric problems other than alcohol or substance abuse is problematic; these subjects can score differently on a variety of neuropsychological instruments. For example, clinically depressed individuals can have reduced performance on effortful neuropsychological tests. A number of investigators have chosen to enroll individuals who have been treated for prior psychiatric problems (McAllister et al., 1992). Other investigators have chosen to exclude these individuals (Martin et al., 1992; Poutiainen et al., 1993; Sinforiani et al., 1991). If such subjects are included in the infected group, then subjects with similar diagnoses should be included in the control group as well. Of course, this will increase variance on test performance, thereby increasing the sample size for sensitive detection of differences.

In addition to the issues of clinically diagnosable

psychiatric problems, even variations at subclinical levels can be an important influence on neurocognitive tests performance. Until relatively recently, investigators believed that HIV+ subjects were more likely to be depressed than control subjects and that depression affects performance. Thus, depression could account for observed between-group differences. However, the majority of the data indicate that between-group differences in depression do not account for differences on the tests (McAllister et al., 1992; Poutiainen et al., 1993). Therefore, depression is only a problem if it is unequally represented in the HIV+ and HIV- groups.

English competency is another problem. The vast majority of neuropsychological and information processing tests have been developed in English-speaking countries. Administering these tests, which involve language, to individuals who have varying degrees of competency in English is problematic. Interestingly, only a few investigators report screening their subjects on the basis of their primary language (e.g., Wilkie et al., 1992).

Sampling. For studies conducted in the United States, three variables--ethnicity, intelligence, and education--have posed serious methodological problems for many studies; all three variables are interrelated and all covary with performance on neuropsychological instruments. In the United States, some minority samples have lower education levels and lower estimated IQs than some White samples. Minority members also tend to have a lower socio-economic status and less access to quality health care, which may make them more willing to volunteer for research

protocols that provide health care. Arguably, lower socio-economic status minority subjects at more advanced stages of the disease may disproportionately volunteer for experimental protocols, resulting in an overrepresentation in the more advanced-stage experimental groups. Because neuropsychological instruments covary with education level, which is in turn correlated with estimated intelligence, purported cognitive declines in infected groups could be attributable to unequal proportions of lower socio-economic minorities in experimental groups, particularly in the advanced-stage groups.

A surprising number of studies that purport to demonstrate decreasing cognitive performance with advancing disease state do not report the percentage of different minorities in each group, making interpretation of the data problematic. However, in several studies some minorities who tend to have lower socio-educational status are disproportionately represented in more advanced disease stages, raising the suspicion that the purported decline in cognitive function with advancing disease may reflect socio-economic factors rather than HIV-related factors. Studies reporting decreasing education levels with increasing severity of disease have similar interpretation problems as do studies reporting a decreasing IQ (Poutiainen et al., 1993).

In a surprising number of studies the control subjects are more educated than the HIV+ subjects (Grant et al., 1987) although the differences are rarely statistically significant (McArthur et al., 1989). Such differences should not be ignored; non-significant

differences still may cause problems in the interpretation of neuropsychological test results. Interestingly, only one study was found in which the HIV+ subjects were more educated than the controls (Sinforiani et al., 1991).

Age is a problem that is rarely addressed in HIV studies. A few studies provide no information on the age of the subjects. Many studies never present the age range of the subjects and do not mention using age-adjusted norms for the neuropsychological measures. Because many neuropsychological measures are affected by age and have published age-adjusted norms, the explicit use and discussion of age-corrected norms should be encouraged. A few investigators (Martin et al., 1992; Wilkie et al., 1992) have elected to use age as a covariate in the statistical analyses of the neuropsychological data.

HIV-specific research issues. An issue for HIV research is that injection drug use is one of the common modes of viral transmission. Some investigators have allowed injection drug users to participate in studies examining the cognitive effects of HIV (McAllister et al., 1992) to provide a comprehensive examination of the effects found with all modes of transmission. Because injection drug use and the commonly co-existing alcohol abuse may result in cognitive deterioration, the HIV+ groups may perform more poorly on cognitive tests because of drug and alcohol effects, not because of HIV per se. However, separating the effects of chronic substance abuse from those of HIV is difficult. Consequently, a number of investigators have tried to exclude injection drug users

as well as alcohol and substance abusers from study participation (Martin et al., 1992; Poutiainen et al., 1993; Sinforiani et al., 1991; Wilkie et al., 1992).

Another HIV-specific issue concerns the definition of "asymptomatic". Using the Centers for Disease Control (CDC) classification system, "asymptomatic" sometimes refers to CDC Stage 2 (Krikorian and Wrobel, 1991; Poutiainen et al., 1993; Rosci et al., 1992). At other times "asymptomatic" refers to CDC Stages 2 and 3 (Bornstein et al., 1992; McAllister et al., 1992). Studies using the Walter Reed Classification System seem to use Stages 1 and 2 as asymptomatic.

Another issue concerns the appropriateness of control groups. Most studies of cognitive processes limit the HIV+ group to homosexual or bisexual males who contracted the disease sexually. Logically, the appropriate control group should consist of HIV- homosexual or bisexual males who are matched to the HIV+ groups on other relevant demographic factors, such as age and education. Interestingly, some investigators have constructed control groups that apparently consist of homosexual, bisexual, and heterosexual males (e.g., Poutiainen et al., 1993). Such control groups may not be comparable to the HIV+ groups on a variety of lifestyle factors that may affect performance on both neuropsychological and information processing instruments.

Enrolling subjects taking anti-retroviral medication is a very important issue; these drugs can affect cognitive processes, reversing the decline in performance for individuals at the more

advanced stages of HIV disease. To date, all three possible strategies for dealing with individuals on anti-retroviral medication have been employed: including these individuals (Wilkie et al., 1992), excluding them (McArthur et al., 1989; Sinforiani et al., 1991), and, in longitudinal studies, initially excluding them but retaining subjects who begin taking the drug during study participation (McAllister et al., 1992). The inclusion of subjects taking anti-retroviral medication has been complicated by the recent approval of two new anti-retroviral drugs, DDC and DDI. Currently, physicians use the three approved anti-retroviral medications (AZT, DDI, and DDC) in various combinations. These combinations may be administered in various cycles, such as receiving DDC one week and DDI the following week. Such combinations pose another problem in the inclusion of subjects taking anti-retroviral medication. Further, additional such drugs are currently available on a compassionate usage basis, complicating the situation even further for future studies.

The final issue concerns the specific tests that are used to assess cognitive decline. Almost by definition, no one study can measure all aspects of human cognitive performance because of limited testing time. Consequently, an investigator must decide either to test a limited number of cognitive domains in depth or to test a larger number of domains superficially. A preliminary review of the literature indicates that most of the studies, particularly large-scale experiments involving repeated measurement of the same individuals, attempt to test a large number of domains

superficially. Because of time constraints, only one or two tests are administered to assess a given domain. If a test is not sensitive to HIV-related deterioration of the particular domain, the results may be misleading. This problem is exacerbated by the fact that a few tests have been employed in HIV assessm batteries repeatedly. The choice of these tests appears to based more on their use in neuropsychological assessment in general and their administration time than on any demonstrated sensitivity to HIV-related impairment.

Use of Reaction Time as a Dependent Measure

Almost all of the early studies of the effects of HIV on cognitive processes used neuropsychological instruments exclusively to assess decline. Neuropsychological tests typically use the percentage or number of correct responses as the dependent measure. Recently, there has been an increased emphasis on tests using very precisely measured reaction time (typically to $\pm$ 1 ms) as the dependent measure. To some extent, this change can be attributed to the work of Martin and Mapou and their colleagues (Mapou, Kay, Rundell, and Temoshok, 1993; Martin et al., 1992). For example, Martin et al. (1992) showed no between-group differences between HIV+ and HIV- subjects on any neuropsychological tests but did find a significant difference on information processing tests that used reaction time as a dependent variable. Additionally, a significantly higher proportion of HIV+ as compared to control subjects had reaction times that were 2 standard deviations (sd)

slower than the mean of the control group on these tests. At each of three six-month follow-up testing sessions, a higher proportion of HIV+ subjects were 2 or more sd below the mean of the controls. Martin et al. (1992) interpreted their results to indicate that a certain proportion of HIV+ individuals suffer cognitive declines relatively early in the course of infection. This interpretation has effectively shifted the controversy from whether or not cognitive decline occurs in early-stage individuals to what percentage of these individuals actually experience a cognitive decline, the nature of the cognitive processes that are affected, and their implications for work performance.

Purpose

The study reported in this document had two primary purposes. The first was to determine the disease stage at which cognitive decrements become detectable in HIV+ individuals. The second was to compare information processing tests, which typically use reaction time as the dependent measure, to neuropsychological assessment.

The second purpose had an operational as well as a general scientific objective. Military personnel frequently must operate in environments where specialized medical personnel are not immediately available. If information processing tests are at least as sensitive as neuropsychological instruments to various types of cognitive impairments, they may be more useful in many military environments because they can be largely automated. This

automation can reduce the training required for the test administrator and facilitates transmission of data to a medical facility.

Approach

To provide some background on the second purpose mentioned above, the section below will give a brief comparison of assessment using neuropsychological versus information processing instruments. Next, some of the specific issues involved in test selection are discussed. Finally, our approach to the methodological issues mentioned earlier are described.

Neuropsychological versus information processing assessment. Neuropsychology and information processing instruments have very different theoretical underpinnings and supporting experiments, and to date, relatively few individuals have made any attempt to integrate or compare the techniques. Each of these types of assessments has specific strengths and weaknesses, which will be discussed briefly below.

Human neuropsychology is basically the study of the neural substrates underlying human behavior. It is a discipline that developed by applying psychological testing principles and methods to neurological syndromes. Thus, neuropsychological testing can be one of the most sensitive measures of brain dysfunction and may be useful in diagnosis of cases where other measures of brain integrity, such as those obtained from brain imaging or electrophysiology, appear unimpaired. Even when there are abnormal

findings from brain imaging studies, the functional significance of these abnormalities can only be determined through behavioral testing. For example, a small lucency in a subcortical structure of the brain can have absolutely no functional significance or can lead to profound changes in cognitive function. The neuropsychologist is called upon to make this determination. In addition, neuropsychological variables are playing an increasingly important role in neurological, neurosurgical, and neuropharmacological research. These respective fields are increasingly recognizing the importance of cognitive disturbances and recognize the value of neuropsychological measures for assessing the outcome of surgery or drug treatment.

Information processing instruments typically have been derived from cognitive psychology, which is an outgrowth of and a reaction to the behavioralistic tradition that dominated psychology from the early years of this century to approximately the late 1960's. Cognitive psychology is concerned with memory, attention, learning, and decision making. Cognitive models rarely refer to anatomic locations in the brain and, for the most part, neurology and neuroanatomy are not considered. Thus, information processing tests developed from cognitive psychology are rarely used clinically to evaluate central nervous system insult. The research in this area typically has been conducted on normal, working-age adults although some studies have been performed on children, the aged, and populations with specific cognitive problems.

Currently, neuropsychological instruments are established

assessment tools in the medical community while information processing tests are almost unknown to physicians. Neuropsychological assessment identifies specific anatomic brain regions with specific tests. Thus, probable sites of lesions or other forms of damage can be identified. Information processing tests generally cannot be linked to any specific region of the brain. Neuropsychological assessments may be conducted on individuals who are seriously impaired. Information processing tests usually cannot be performed by individuals with major cognitive deficits. Neuropsychological instruments generally have not been related to complex tasks that may be encountered in certain demanding jobs. In contrast, some information processing tests have been used to predict performance on complex tasks, such as flying. Finally, neuropsychological instruments generally cannot be administered repeatedly during short time intervals without compromising the validity of the tests. In contrast, information processing tests may be used in a repeated-measures testing protocol with little concern for the adverse affects of practice. The use of a repeated measures protocol also allows between-group differences in learning to be examined. Such information can provide valuable insights into relatively subtle task effects.

Selection of information processing tasks. When this study began, information processing tests had not been used to examine the effects of HIV on cognitive processes. Thus, it was not possible to select information processing tests that were demonstrably sensitive to the effects of HIV. Instead, the

investigators decided to develop a preliminary battery and determine the sensitivity of the instruments comprising the battery during a pretest.

The initial task selection was based on two criteria: patient complaints and guidance from the program manager to select tasks with known relations to naval jobs. The project physician was questioned concerning common patient complaints about cognitive functioning. Many patients complained about a deterioration in what appeared to be spatial processing. Consequently, a test of spatial processing was included in the initial pretest. The specific test selected, the matrix test, had been evaluated as a potential aircrew selection test. Many patients also complained about their inability to concentrate. Thus, a test of sustained attention was included in the preliminary test battery. This test, a classical vigilance test, has been used for many years as a laboratory simulation of radar and sonar operators' jobs. A third test, the running difference test, was included because it previously had been evaluated for use in the naval aviator selection battery. Finally, the Sternberg Memory Search Task (Sternberg, 1969) was included because preliminary data from other naval investigators indicated that it was not affected by the progress of HIV. Thus, it was included to act as a type of control.

The subsequent pretest found that all of the tasks except the Sternberg Task showed performance decrements, indicating possible sensitivity to HIV-induced cognitive decrements. The Sternberg Task

(Sternberg, 1969) was retained in the battery because, as noted above, it acted as a type of control task, and because it is one of the few information processing tasks that has normative values. Additionally, two derived measures can be obtained from this task that theoretically measure different cognitive processes and are known to be affected by factors such as alcohol consumption and exposure to mercury. Thus, one test can be used to examine the effects of HIV on several different cognitive processes.

The dependent measures for all four tasks have good characteristics for statistical analyses and for repeated administration: high intertrial correlations with consistent standard deviations. The primary dependent measure for all of the tasks except vigilance was reaction time.

Selection of neuropsychological instruments. The neuropsychological tests were selected using several criteria. First, they measured neuropsychological domains that had been found to be sensitive to deficits in HIV+ individuals. Second, the tests assessed a wide range of neuropsychological domains to permit examination of the intercorrelations among the information processing and the neuropsychological tests. Third, because of a 3-hour time limit, tests that might be sensitive but require an extended time to administer (e.g., the Halstead Category Test) were omitted. One exception to this was the Paced Auditory Serial Addition Test (PASAT), which was included because previous research had found that asymptomatic individuals performed more poorly on this test than control subjects (Grant et al., 1987). Fourth, some

of the tests appeared to assess the same functions as the information processing tests.

In addition to being sensitive to HIV-related deficits, a subset of tests was selected to provide estimates of premorbid abilities. Clearly, we did not want to attribute differences between the HIV+ and the HIV- subjects to the disease process when these differences actually reflected pre-infection differences. The National Adult Reading Test (NART-R) and the Vocabulary and Information subtests of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) were included to estimate premorbid abilities.

Methodological Issues

Twelve major methodological issues were mentioned earlier. We will discuss our approach to controlling 11 of these. The twelfth, test selection, was discussed earlier. The most difficult issue was that of substance abuse. We did not enroll individuals who had participated in any alcohol or drug treatment programs or who reported any injection drug use. We also established limits for the frequency and quantity of substance use, which were rigorously observed. Nevertheless, as discussed earlier, self report is problematic for substance abusers. Consequently, we questioned the subjects repeatedly at different points in the protocol concerning their substance use. This allowed us to check the consistency of their answers. Further, a toxicological screen was added at the time of the physical examination, and subjects were told that the screening would occur.

The same general approach was used for psychiatric problems: Any candidate reporting treatment for psychiatric problems was excluded. Again, subjects were questioned repeatedly about psychiatric problems to check their consistency of response.

The issues concerning ethnicity, intelligence, age, and education were addressed in several different ways. First, the proportion of minorities in each experimental group was kept constant and groups were matched on education. Matching for education also eliminated the problem of more educated control groups. Second, as noted earlier, several estimates of pre-morbid intelligence were obtained to allow for statistical corrections, if necessary. Third, to eliminate the problems associated with low English fluency, only native English speakers or individuals acquiring English by age 6 were enrolled. Fourth, age effects were controlled by restricting the age range of the subjects to ages 21 to 50.

To account statistically for any between-group differences in depression, we administered several mood scales to the subjects as part of the neuropsychological battery. Scores on these scales could be used as covariates if between-group differences in depression were found.

Although there are several modes of HIV transmission, we decided to enroll only individuals who had contracted HIV through sexual transmission for two reasons. First, as noted earlier, injection drug use can have adverse effects on cognitive processes, confounding the effects of substance abuse with those of HIV.

Second, informal discussions with the staff at the Los Angeles County/University of Southern California medical facilities indicated that few potential subjects were available who had contracted the disease through blood products. Thus, we decided to concentrate on the most common route of infection, sexual transmission.

Two groups of control subjects were included in the study. One group consisted of HIV-, homosexual males. These subjects were included to control for lifestyle differences that may affect performance on various assessment instruments. A heterosexual control group was included to allow comparisons to military personnel.

Initially, we did not enroll anyone taking anti-retroviral medication. Because anti-retroviral medication can have cognitive effects, we wanted to exclude individuals whose performance on both the information processing and the neuropsychological batteries could be attributable both to the effects of HIV and to the medication. However, as the study progressed, the use of anti-retroviral medication became increasingly common in symptomatic individuals. Consequently, a fifth group composed of subjects taking anti-retroviral medication was added to the study to assess directly the effects of such medications on performance.

METHODS

Design

Subjects were recruited into five groups. The first group consisted of asymptomatic HIV+ males in Walter Reed Stages 1, 2, and 3. The second group consisted of symptomatic HIV+ males in Walter Reed Stages 4 and 5. The third group consisted of HIV+ males who were taking AZT, DDI, DDC, or any combination of these drugs. The design also included two control groups of HIV- males. The fourth group consisted of homosexual males; the fifth group, of heterosexual males. The five groups were matched on the distribution of scores for age and education (including the mean and range). The proportion of minorities in each group was kept approximately equal.

Preliminary calculations of statistical power indicated that 28 subjects would be required in each group.

Subjects

Recruitment. Both HIV+ and HIV- volunteers responded to newspaper advertisements and notices placed in local colleges. In addition, the HIV+ volunteers responded to notices distributed through Los Angeles area HIV support groups, Los Angeles HIV clinics, individual physicians, the Los Angeles County/ University of Southern California Medical Center, and a variety of newsletter advertisements.

Initial telephone screening. Candidates completed a 20 to 25 min screening interview to determine their eligibility for participation. The inclusion criteria were: age between 21 and 50 years, at least 12 years of formal education (including the receipt of a high school diploma), and English as the native language or acquired by age 6. The exclusion criteria included a self-reported history of: treatment for alcoholism or drug abuse, head injury or other episodes with a loss of consciousness greater than 5 min, convulsions or seizures, treatment for severe mood swings or other psychotic disorders, stroke, brain tumor, cancer (other than skin cancer), abnormal blood pressure, heart disease, diabetes, an AIDS-defining illness, and uncorrected sensory deficits that would interfere with understanding the test instructions or performing the tests. Additionally, anyone participating in protocols that used any of the same neuropsychological tests was excluded unless the data could be shared between the two studies. However, only two people completed the study who were enrolled in other protocols. One of the subjects was in the anti-retroviral medication group; the other, in the asymptomatic group.

Use of psychoactive substances also was an exclusionary criteria. Respondents were questioned about their use of over-the-counter medications, prescriptive medications, recreational drugs, and alcohol. Respondents who used medications with psychoactive effects on an ongoing basis were excluded unless the medication was prescribed for treatment of HIV infection (e.g., AZT) and was

obtained from a licensed source. Respondents were excluded if their alcohol use was greater than 12 drinks per occasion or if their alcohol use exceeded 6 drinks and their frequency was greater than two times per week. A drink was considered to be 12 oz. of beer, 4 oz. of wine, or 1.5 oz. of liquor. Respondents reporting use of the following drugs with a frequency greater than once per week were also excluded: marijuana, stimulants, tranquilizers, sleeping medications, amyl nitrates, or steroids. Occasional use of any of these drugs did not exclude participation, however.

If the subject successfully completed the telephone screening, he was conditionally accepted into the study and scheduled for a physical examination. Participants were notified at this time that a toxicological screening would be done during the physical examination. All subjects conditionally accepted into the study agreed to refrain from drinking alcohol for 48 hours before the physical examination and to refrain from using any recreational drugs for 1 week prior to the physical examination. Subjects who completed the entire study were paid $70.00 for their participation.

Physical, neurological, and toxicological screening. Secondary screenings were conducted during the physical examination and involved testing for illnesses that could influence performance on the cognitive tasks, for neurological disorders that were clearly not related to HIV, and for recreational drugs. Control subjects were screened as rigorously as HIV+ subjects.

The reasons for excluding enrolled subjects from the study are

summarized in Figure 1 and discussed throughout this manuscript. One subject failed to report for the physical examination. One subject was found to have an AIDS-defining illness during the examination and was excluded from continued participation. Five subjects, all from the homosexual control group, reported flu-like symptoms and had abnormal blood chemistry panels. These subjects were also excluded. Five other subjects with clinically significant medical problems were identified and excluded from the study. Two of these subjects were in the heterosexual control group and three were in the homosexual control group. No subjects were excluded for neurological problems.

The data from 20 subjects were discarded because of positive results on the toxicological screening. Of these, three were in the asymptomatic group, five were in the symptomatic group, eight were in the homosexual control group, two were in the heterosexual control group, and two were in the group receiving anti-retroviral medication.

Additional toxicological, health, and cognitive screening. Additional screenings took place when the subject was seen for the information processing and neuropsychological batteries. The first screening was administered at the beginning of the information processing battery when the subject was asked about drug use in the preceding 30 days. Three subjects--two from the homosexual control group and one from the asymptomatic group--disclosed that they were taking psychoactive drugs and were subsequently dismissed from the study. One subject in the homosexual control group was

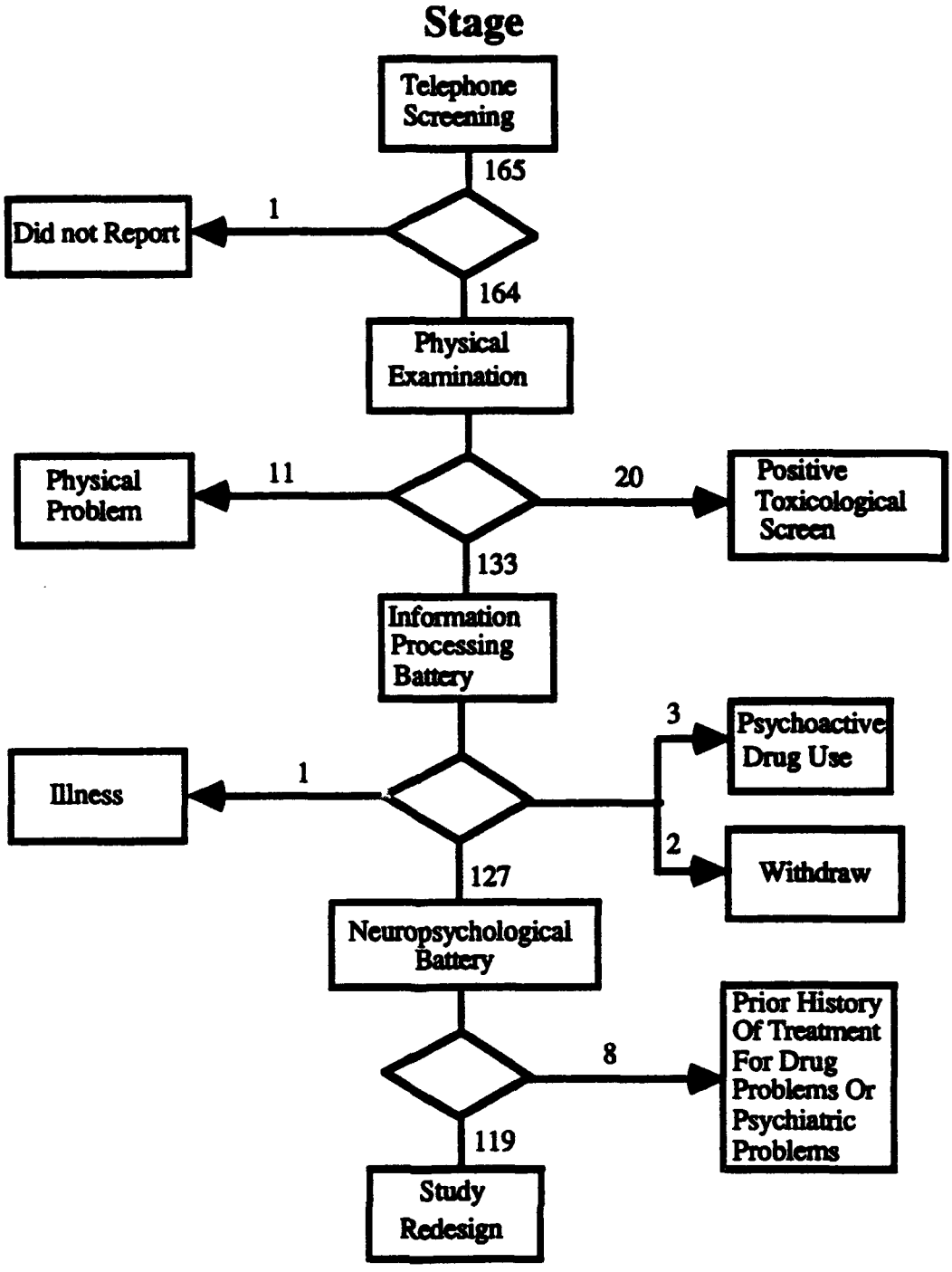

Figure 1. *Reasons for elimination from study. The numbers above the horizontal lines indicate the number of subjects dropped for the reason given in square shaped boxes. The numbers next to the vertical lines show the number of subjects continuing to the stage.*

clearly ill and was also dismissed. The second screening involved testing the subject's reading level using the Wide Range Achievement Test-Revised (Jastak and Wilkinson, 1984). A seventh-grade reading level was required. No subjects were excluded for having less than a seventh-grade reading level.

Two additional subjects were lost during this phase of the study. One of the heterosexual control subjects decided to withdraw from the study during the information processing battery. One of the homosexual control subjects completed the information processing battery but failed to return from a lunch break for the neuropsychological battery.

During the neuropsychological testing, the examiner followed up on any unusual responses to the subjective mood scales or the standardized tests. During this questioning, eight subjects revealed histories of psychiatric problems or drug abuse that had not been reported during the initial telephone screening. These subjects were dismissed from the study. Two of these subjects were in the asymptomatic group; two, in the symptomatic group; two, in the homosexual control group; one, in the heterosexual control group; and one, in the anti-retroviral medication group.

Apparatus for the Information Processing Battery

A Hewlett-Packard QS/16S microcomputer generated all stimuli, recorded and processed the subjects' responses, and timed all trials. All responses were made on a modified Texas Instruments 99/4 alphanumeric keyboard placed on the same table as the

computer. The centermost row of keys was used for responses. All instructions and stimuli were displayed on a Hewlett-Packard color VGA monitor. All alphanumeric stimuli were 2.5 cm in height. The subjects sat 66 cm from the monitor.

Information Processing Battery

The order and duration of the tasks comprising the information processing battery are given in Table 1 with approximate administration times. The Sternberg Memory Search Task, the matrix task, and the running difference tasks all had 90-s trials separated by a 30-s break. Feedback consisting of the number of correct responses, the percentage of correct responses, and the average correct reaction time was presented during the breaks.

Two dependent variables, the average correct reaction time and the percentage of correct responses, were recorded for these three tasks. Two dependent variables were recorded for the vigilance task: hits and false alarms.

Sternberg Memory Search Task. The Sternberg Memory Search Task measures the speed at which a displayed stimulus is compared to a representation in memory (Sternberg, 1969). At the beginning of each trial, the subject saw a randomly generated positive set of letters, which he memorized. Each letter was presented for 1 s. A tone sounded and the task began 1.5 s after the presentation of the last member of the positive set. The subject then saw a series of randomly generated letters, approximately 50% of which were members of the positive set. The subject pressed the key (either F or J)

Table 1

Order of Tests and Approximate Administration Time

Test	Approximate Administration Time In Minutes
Information Processing	
Drug Use Questionnaire	NA
WRAT	3
Handedness Inventory	2
Sternberg	65
Matrix	45
Vigilance	48
Running Difference	45
Neuropsychological Battery	
Subjective Memory Questionnaire	NA
POMS	NA
Beck	NA
Finger Tapping	8
CVLT, Immed. and Short term	25
Complex Figure Copy and Immed.	7
WAIS-R Digit Span	5
WAIS-R Digit Symbol	3
Grip Strength	4
Grooved Pegboard	6
CVLT 20 Min. Delayed recall	5
Complex Figure Copy Delayed Recall	5
WAIS-R Block Design	10
Stroop	7
Repeatable Episodic Memory	6
WAIS-R Picture Arrangement	10
Wisconsin Card Sort	15
PASAT	15
NART-R	5
Oral Fluency	6
Trail Making Part A	2
Trail Making Part B	3
WAIS-R Vocabulary	12
WAIS-R Information	8
Boston Naming Test	7

under the index finger of his dominant hand if the letter was a member of the positive set and pressed the key under the index finger of his non-dominant hand if the letter was not a member of the positive set. After successfully completing one trial atpositive set size 2, the subject completed four additional trials at set size 2, followed by five trials at set size 3, 4, 2, 3, and 4 for a total of 10 trials per set size. At the end of each trial, the subject received feedback on his performance.

The Sternberg task was the only test in the information processing battery that did not describe performance using raw data. The average correct reaction time and the average percentage of correct responses were calculated for each set size and for each response type (*Yes* versus *No*). The average correct reaction times for *Yes* were then regressed on positive set size using a linear equation, and the slope and the intercept of the linear equation became the dependent measures of interest. This process was repeated for *No* responses. Thus, three dependent measures--slope, intercept, and accuracy--at two levels of the response type variable, *Yes* versus *No*, were calculated for this task.

Matrix. The second task assessed spatial short-term memory. In this task 5 X 5 matrix grids were presented sequentially to the subject. Each matrix had five, randomly selected, illuminated cells. The subject's task was to determine as quickly as possible if the current matrix was identical to the preceding matrix. If the current matrix was identical to the preceding matrix, it was not presented in the same orientation; the second of the two identical

matrices was always rotated 90° relative to the first matrix. If the current matrix was a rotated version of the immediately preceding matrix, the subject responded *same* by pressing the key under the index finger of his dominant hand. If the current matrix was different, he pressed the key under the index finger of his non-dominant hand. The response to the first matrix pattern of any trial was always *same*. Approximately 50% of the correct responses on any given trial were *same* and 50% were *different*. A pattern was allowed to repeat itself a maximum of four times. The matrix was a 11.8 cm. square.

After successfully completing one trial, subjects were given three blocks of 4, 5, and 5 trials, respectively.

Vigilance. This task was the primary measure of sustained attention in the information processing battery. It was a version of the classic tests used to measure sustained attention over long (approximately 1 hr) periods. For this task a pair of 0.5 cm dots was illuminated at the center of the monitor. These dots were 10 cm apart and were displayed for 150 ms each s. The target stimuli consisted of the same pair of dots illuminated for the same interval. However, the dots now were 12.5 cm apart. The subject pressed the key (either F or J) under the index finger of his dominant hand when the target was identified. The subject had to correctly identify 80% of the targets presented during a 60-s training period with no more than two false alarms to begin the test. If the subject failed to reach criteria, he repeated the training period until he met the criteria. The test required 48 min

and had an average interstimulus interval of 89 s and an interstimulus range of 26 to 192 s. At the end of the testing session, the subject was presented with the percentage of correctly detected signals, the number of false alarms, and his average correct reaction time to the signals.

Running Difference. This task assessed verbal short-term memory and simple arithmetic skills that are important in many daily activities. In this task a series of digits ranging from 0 to 8 were presented. The subject's task was to determine the difference between the currently presented digit and the immediately preceding digit. The subject used the keys under the index through little finger of his dominant hand to enter 1, 2, 3, and 4, respectively, whereas the keys under his index through little finger of the non-dominant hand were used to enter 5, 6, 7, and 8, respectively. Two dependant variables, the percentage of correct responses and the correct response time, were recorded. After successfully completing one trial, subjects were given three blocks of 4, 5, and 5 trials, respectively.

Neuropsychological Battery

The tests and the subjective mood scales used in the neuropsychological battery are listed in Table 1 in the order in which they were presented. The approximate completion times also are given. The neuropsychological tests produced a total of 77 variables; the subjective scales produced eight. All but five of the tests have been described by Lezak (1983) and will not be

elaborated here. These five are the National Adult Reading Test (Nelson and O'Connell, 1978), the California Verbal Learning Test (Delis, Kramer, Kaplan, and Ober, 1987), the University of Southern California Repeatable Episodic Memory Test (Parker, Bridge, Ingraham, Eaton, and Heseltine, 1989), the Profile of Mood States (McNair, Lorr, and Droppleman, 1971), and the Subjective Memory Questionnaire (Squire and Zouzounis, 1988).

National Adult Reading Test (NART-R). The National Adult Reading Test-Revised (Blair and Spreen,1989) consists of 61 words that can not be phonetically decoded correctly. The test was administered according to instructions for the National Adult Reading Test (Nelson, 1982). The number of errors was recorded and an estimated premorbid IQ was derived from the following equation: IQ = (-.78) (# of errors)+ 127.8 (Nelson, 1982).

California Verbal Learning Test (CVLT). Subjects were presented with a list (List A) of 16 words from four general categories--tools, clothing, spices and herbs, and fruits--at the rate of one word per s. Following free recall, subjects were given four additional presentation/recall trials. Subjects were then presented with a second list (List B) of 16 words also from four general categories--utensil, fish, spices and herbs, and fruits--at the rate of one per s. Following List B recall, subjects were asked to recall items from List A (short delay free recall). Subjects were then given a cued recall trial for List A items (short delay cued recall). The general categories were given and subjects were asked to recall items from that category. Following a 20 to 25 min

delay (filled with other non-related tests), subjects were given a second free recall (long delay free recall) and cued recall (long delay cued recall) trial for items in List A. The administration of the CVLT ended with a 44 item recognition test (long delay recognition) that included all 16 items from List A, 8 items from List B, and 20 distracter items.

A number of measures relating to the strategies and processes involved in learning and remembering verbal material can be derived from the CVLT. Recall measures include total recall from List A, recall from List B, and short- and long-delay free and cued recall. Learning measures include the extent to which subjects clustered items semantically or according to serial position, the increment in words recalled per trial, and the consistency of recall across trials. Types of recall errors include perseverations and intrusions. Recognition measures include correct recognition hits, discriminability, and the number of false positives.

University of Southern California Repeatable Episodic Memory Test (REMT). Subjects were presented with a list of 15 unrelated words at a rate of one every 2 s. Following a 1 min period for free recall, subjects were presented with two additional presentation/recall trials. The items, presentation rate, and recall rate were identical to those of Trial 1; however, the order in which the items were presented was altered for both Trial 2 and Trial 3. The number of items correctly recalled, repetitions, and intrusions were tallied for each trial and totals were summed across trials.

Profile of Mood States (POMS). The POMS is a 65-item adjective rating scale reflecting current mood states (McNair, Lorr, and Droppleman, 1971). Each adjective is rated on a 5-point scale from 0 (not at all) to 4 (extremely) according to how the subject had been "feeling during the past week, including today". The test was administered and scored according to standardized procedures. Scores were obtained for each subject on the six POMS factors: tension/anxiety, depression/dejection, anger/hostility, vigor/activity, fatigue/inertia, and confusion/bewilderment.

Subjective memory questionnaire. This is an 18-item self-rating of memory abilities (Squire and Zouzounis, 1988). Ratings were made on a 9-point scale ranging from -4 (worse than ever before) through 0 (same as before) to +4 (better than ever before). Subjects rated their ability to remember items now compared to the time period before the last year. Examples are: "My ability to recall things when I really try is ...", "My ability to recall things that happened a long time ago is ...", and "My ability to remember names and faces of people I meet is ...". A total score was computed for each subject and could range from -72 to +72.

Procedure

History and physical examination. Each subject underwent a complete medical history with emphasis on eliciting symptoms attributable to AIDS dementia complex, as well as a drug and alcohol history. A physical examination, including a careful neurologic examination, was also performed on all subjects, again

with special emphasis on HIV-related findings (e.g., oral hairy leukoplakia). A complete blood count and chemistry panel were performed as well as a urine analysis to determine compliance with alcohol and recreational drug use restrictions.

Each subject was tested for HIV seropositivity, employing the ELISA antibody screen. P24 antigen levels in serum were also evaluated. In addition, T-cell subsets on peripheral blood were analyzed, as well as the absolute numbers of T4 and T8 lymphocytes. Delayed cutaneous hypersensitivity testing was performed using the Merieux skin test panel, a battery consisting of seven recall antigens.

In addition, the Karnofsky Performance Score (Karnofsky and Burchenal, 1949) and the Mini-Mental State test (Folstein, Folstein, and McHugh, 1975) were administered. At the start of the examination, the subject completed a Symptoms Questionnaire that was designed to elicit symptoms specifically associated with HIV infection. The entire history and physical, mental status, and blood examinations required between 1.5 to 2 hr.

Information processing battery. The subjects performed the information processing and neuropsychological batteries 2 days after the physical examination. After completing the drug and reading level screens, the subjects completed the Edinburgh Handedness Inventory (Oldfield, 1971).

Subjects then began the information processing battery. The procedures for the Sternberg Memory Search Task, the matrix task, and the running difference task were identical. After the subject

finished reading the instructions, he was given an opportunity to ask questions. When all questions were answered, the subject was asked to describe the test procedure. If the description was adequate, he performed one trial. If the percentage of correct responses was low, then the instructions were recalled and discussed until the experimenter felt that they were understood. If the percentage of correct responses was high, the subject began the task. The experimenter avoided speed/accuracy instructions unless the subject asked specifically about this trade-off. When these questions occurred, the investigator told the subject to attend to both aspects of his performance as well as possible. Subjects were given approximately a 1-min break as the computer was reset after each block of trials as well as a 10 to 15 min break between the second and third tasks. The total time for this section of the testing was approximately 3.75 hr.

Neuropsychological test battery. The neuropsychological test battery was administered following a 30 to 60 min break after completion of the information processing battery (for those subjects completing both batteries in one day). The neuropsychological battery took approximately 2.5 to 3.5 hr to administer. Just before the subject began the battery, his Merieux skin test was read.

RESULTS

Approach

The original proposal stated that the asymptomatic group would be defined by Walter Reed Stages 1, 2, and 3. After the data were collected, research at the Naval Medical Research Institute indicated that the standard method of administering the Merieux test, which was used in this study, produced unreliable results. The unreliability of this measure called into question the subject classifications in general and the Stage 3 versus Stage 4 assignments in specific. Because the distinction between a subject at Walter Reed Stage 3 versus one at Stage 4 rests on the results of the Merieux test, the original design had to be discarded.

In consultation with physicians at the Naval Health Research Center, we developed another experimental design that divided the HIV+ subjects into three groups. The first group, the asymptomatic group, consisted of subjects at Walter Reed Stages 1 or 2. The second group consisted of subjects at Walter Reed Stages 3 or 4. The third group was composed exclusively of Walter Reed Stage 5 subjects. Classifying the subjects into these three groups does not require reliable results from the Merieux; this classification can be done using the results of the history and the blood panels alone.

Towards the end of the data collection effort, we realized that community physicians had begun administering AZT to asymptomatic subjects as a means to delay HIV disease progression;

AZT was licensed for this purpose. We had originally assumed that all of the subjects in the anti-retroviral group would be symptomatic (i.e., at Walter Reed Stages 3, 4, or 5). However, by the end of the data collection period, we realized that this assumption was incorrect. We had decided to divide the anti-retroviral subjects into two groups based on their Walter Reed Stage when we discovered the problem with the Merieux Test. While we were developing the new design, we realized that the subjects taking anti-retroviral medication probably were misclassified according to standard clinical practice; individuals typically are placed on anti-retroviral medication after their T4 cell count falls below 500 and/or they display certain symptoms. Neither the Walter Reed nor the CDC system allow individuals to regress to a less symptomatic stage despite changes in the person's apparent health. Thus, our classification of the subjects taking anti-retroviral medication probably was biased towards the asymptomatic stage because the anti-retroviral medication typically raises the T4 cell count for a short time after initial use and is associated with a delay in the progression of the HIV disease. Because we had no way to determine the subject's true stage prior to receiving the anti-retroviral medication, we decided to omit all of the subjects taking anti-retroviral medication from the analyses.

Before analyzing the data, we had to examine all of the subjects to ensure that they were placed in the correct stage given the unreliability of the Merieux. Three subjects no longer

could be classified. The subjects who were originally in Walter Reed Stage 5 were examined particularly carefully; placement in this stage could rest predominately on the results of the Merieux or on symptoms.

After these adjustments, we had three subjects in Stage 3, five in Stage 4, and two in Stage 5. With five or fewer subjects in any of the symptomatic Walter Reed stages (Stages 3, 4, and 5) and only a total of ten subjects in all three stages, an analysis of variance (ANOVA) design including a symptomatic subject group would seriously compromise the statistical power of the test; power depends largely on the smallest cell size in the design. Additionally, we were doubtful of the meaning of an analysis combining subjects from the three stages. In discussions with Navy physicians at the Naval Health Research Center our concerns about the interpretability of such an analysis were confirmed. Thus, with only ten confirmed symptomatic subjects, we decided to omit the symptomatic group from the analysis and focus on comparisons between asymptomatic subjects and the controls. The subsequent analyses are based on comparisons of the homosexual control subjects (N=29), the heterosexual control subjects (N=28), and the asymptomatic HIV+ group (Walter Reed Stages 1 and 2) (N=29). The redesign of the study and the corresponding loss of subjects is shown in Figure 2.

Comparison of the Two Control Groups

The two control groups are compared on a number of demographic

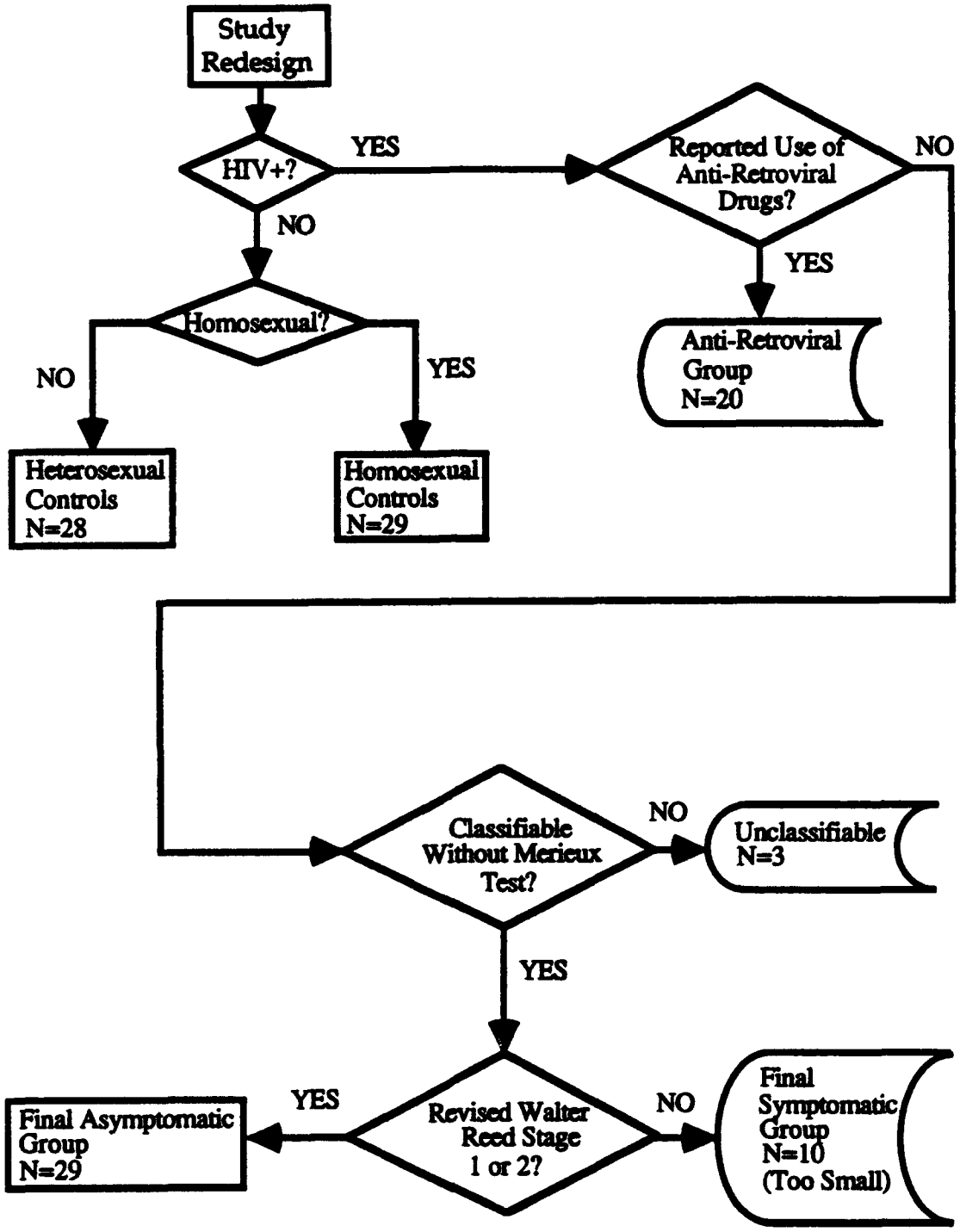

Figure 2. Flow Chart Showing Final Allocation of Subjects to Groups.

variables in Table 2. We performed *t*-tests on the 77 variables from the neuropsychological test battery. Nonsignificant results for these analyses and all subsequent analyses are those for which $p>.05$. Significant differences were found between the two control groups on the first 14 variables listed in Table 3.

A two-way (response type by group) analysis of variance (ANOVA) was conducted on the slope and intercept of the Sternberg Task; a four-way (response type by group by set size by trial) ANOVA was conducted on the accuracy scores. The main effect of group was significant for the intercept of the Sternberg Task [$F(1, 55) = 8.08$, $p=.006$]. The ANOVA conducted on the accuracy measure showed a significant Group X Response Type interaction [$F(1, 55) = 8.73$, $p=.0046$]. See Table 3.

The subjects were also compared on six information processing scores and eight measures of subjective mood. All of these tests were *t*-tests. The tests conducted on the other information processing scores were all nonsignificant except for the average correct reaction time for the running difference task. The groups also differed on the Vigor Scale of the POMS. These differences are also given in Table 3.

These analyses suggest that lifestyle differences are related to differences in cognitive function. To control for these factors, we omitted the heterosexual control group from all subsequent analyses and used only the homosexual control group in our comparisons.

Table 2

Demographic Variables for the Two Control Groups

	Group	
	Heterosexual	Homosexual
N	29	28
Age	34.5 (6.1)	34.8 (7.2)
Education Class (median)	3	3
Ethnicity		
White (%)	71.4	79.3
Other (%)	28.6	20.7
Average Frequency of Drinking (Occasions/Month)	4.26 (3.7)	3.75 (4.4)
Alcohol (Drinks/Episode)	1.83 (1.1)	1.71 (1.9)
NART IQ	113.2 (6.3)	113.1 (8.8)
WAIS VOC	13.2 (2.7)	12.5 (2.1)
WAIS INFO	12.6 (2.2)	12.6 (2.2)
Drug history (% using in the last month)	7.1	34.5

Standard deviations appears in parentheses

Table 3

Means and Standard Deviations for Variables Showing Significant Differences Between the Control Groups

	Group			
	Homosexual Control		Heterosexual Control	
Measure	mean	s.d	mean	s.d
California Verbal Learning Test				
Total Recall	53.00	(6.23)	57.96	(6.94)
List A Trial	12.21	(1.63)	13.68	(1.63)
List A short-delay free recall	11.10	(2.51)	12.50	(2.59)
List A short-delay cued recall	11.79	(2.21)	13.29	(2.02)
List A long-delay free recall	11.38	(2.48)	12.93	(2.49)
List A long-delay cued recall	11.86	(2.60)	13.64	(1.87)
Percent Recalled	80.79	(6.68)	85.82	(5.72)
Recognition Hits	14.34	(1.61)	15.11	(0.83)
Discriminability	93.69	(4.53)	96.82	(3.09)
False Positives	1.10	(1.29)	0.46	(0.79)
PASAT				
Series 1 No. Correct	37.55	(10.58)	44.89	(6.57)
Series 2 No. Correct	32.83	(11.01)	39.29	(8.56)
Series 3 No. Correct	29.11	(10.00)	35.04	(6.24)
Grip Strength				
Dominant Hand	48.07	(6.77)	52.54	(8.15)
Running Difference Task				
Average Correct RT	2067.12	(436.94)	1674.66	(490.30)
Sternberg Task				
Intercept	434.8	(89.79)	376.0	(76.21)
Accuracy (*Yes*)	96.75	(2.62)	93.57	(7.82)
Accuracy (*No*)	98.15	(2.00)	98.32	(2.21)
POMS				
Vigor	63.34	(10.92)	56.96	(6.74)

Group Comparisons on Demographic, Intelligence, Recreational Drug Use, and Depression

Neuropsychological variables are known to be affected by a number of variables, such as age, intelligence, and substance abuse. They also may be affected by depression. Consequently, it is important to demonstrate that any groups that are to be compared using neuropsychological instruments are not significantly different on variables that might affect performance. We categorized the variables that could affect performance into four major groups: demographic variables, intelligence, recreational drug usage, and depression. Under the demographic group, three variables were examined: age, education, and ethnicity. Under the intelligence category, the WAIS-R Vocabulary, the WAIS-R Information, and the NART-R estimated IQ were examined. Depression was examined by the Beck Depression Inventory and the POMS Depression Scale. Finally, as estimates of substance abuse, the number of drinks/month, the average number of drinks/occasion, and admission of recreational drug use in the last month were examined. The means and standard deviations for both groups on these four classes of variables are given in Table 4.

T-tests were calculated for all of the variables mentioned above except education class, ethnicity, and admission of drug use. None of the *t*-tests showed significant between-group differences ($p>.05$) (See Rogers, Howard, and Vessey, 1993 for a different approach to testing group equivalences.)

Determining the number of full-time years of formal education for subjects with advanced degrees was sometimes difficult. Consequently, the subjects were placed into one of five education classes--high school only, some college, bachelor's degree, some post graduate education, and a professional degree, such as a Ph.D or an M.D. Because of the small number of subjects in the professional degree category, the two highest categories had to be combined before any analyses were conducted. A χ^2 analysis with four groups showed no significant between-group differences. However, because of the small number of subjects in some of the cells, this analysis was repeated combining the two lowest education categories, high school only and some college. Again, the three-group analysis failed to reveal any significant between-group differences.

Between-group differences in ethnicity also were examined. Because of the small number of non-White subjects (six in each group), all non Whites were placed into one group. A χ^2 revealed no significant between-group differences in the observed proportion of non-White subjects between the controls and the asymptomatic subjects.The subject's drug history in the last month also was examined using a χ^2 analysis. This analysis (group X used/abstained) also revealed no probabilistic dependency between drug history and group membership.

Table 4

Characteristics of the Two Groups on Demographic Variables, Estimates of Premorbid Function, Mood, and Substance Use

	Group	
	Control	WR 1 and 2
Age	34.52 (6.1)	33.28 (5.5)
Educational Class (median)	3	3
Ethnicity		
White (%)	79.3	79.3
Other (%)	20.7	20.7
WAIS-R VOC	13.21 (2.7)	12.45 (3.3)
WAIS-R INF	12.55 (2.2)	11.59 (3.0)
NART IQ	113.21 (6.3)	110.17 (9.0)
BECK	5.72 (6.2)	9.14 (6.9)
POMS		
Tension	41.24 (8.8)	41.21 (8.4)
Depression	41.62 (8.3)	42.79 (8.7)
Anger	46.48 (10.0)	45.52 (7.2)
Vigor	63.35 (10.9)	60.45 (11.4)
Fatigue	43.76 (7.9)	45.86 (8.6)
Confusion	40.72 (8.4)	42.86 (9.2)
Ave # of drinks/episode	1.83 (1.1)	1.29 (1.0)
Ave frequency of drinking (drinks/month)	4.26 (3.7)	3.71 (4.4)
Drug history (% using in the last month)	34.5	24.1

Standard deviations appear in parentheses

Group Comparisons on the Neuropsychological Measures Using *Z*-Scores

Given what is known about neuropsychological deficits in patients with HIV infection, we would not expect all of the 77 dependent variables of the neuropsychological battery to be equally sensitive. Consequently, the project neuropsychologist designed the focused analysis. The first step was to select a priori a subset of variables that measured different aspects of cognitive functioning and might be sensitive to HIV-related neuropsychological impairments. As described in the Methods and shown in Table 1, the neuropsychological test battery consisted of 15 neuropsychological tests. These 15 tests were initially selected because previous research suggested they might be sensitive to decrements in HIV+ individuals. These tests were: 1) CVLT, 2) REMT, 3) WAIS-R Digit Span, 4) WAIS-R Block Design, 5) WAIS-R Digit Symbol, 6) Trail Making Part B, 7) Complex Figure, 8) Stroop, 9) Wisconsin Card Sorting, 10) PASAT, 11) Oral Fluency, 12) Boston Naming, 13) Finger Tapping, 14) Grooved Pegboard, and 15) Grip Strength. Three additional tests were administered as possible estimates of premorbid functioning. These were the NART-R, WAIS-R Vocabulary, and WAIS-R Information. Two other tests were administered in case they might be needed for a more thorough analysis, but they were not considered essential for the detection of HIV-related decrements and were not included in the focused analyses. These were Trail Making Part A, which would be examined if the results from Trail Making Part B were significant, and WAIS-R Picture Arrangement, which would be examined if critical analyses

pointed to further assessment of frontal lobe structures.

Most of the 15 different tests yielded more than one dependent variable. As a first step, the neuropsychologist selected a limited set of target dependent variables. In some cases the target dependent variable was the one that is routinely used and, therefore, it is the measure for which norms are available. For example, in the case of Trail Making Part B the most common measure is Time To Complete, so it was selected over the variable, Number of Errors. In some cases norms are available for multiple dependent variables of a test. In that case the target variable or variables for a test were those that were most likely to be sensitive to HIV-related impairments. For example, in the case of free recall tests of episodic memory, a number of studies have found that consistency of recall across trials is decreased in HIV-infected individuals. For this reason, recall consistency was selected from among more than ten different variables on the CVLT. Accordingly, 19 dependent variables were selected from the 15 different tests.

For the next step in the analyses each of the 19 variables was transformed to a standard score based upon external norms. The transformation for six of the tests--Oral Fluency, CVLT, Trail Making Part B, Finger Tapping, Grooved Pegboard, and Grip Strength--used published sample estimates of population distributions from the literature. For the other nine tests (Boston Naming, WAIS-R Digit Symbol, WAIS-R Digit Span, WAIS-R Block Design, Complex Figure, REMT, PASAT, Wisconsin Card Sort, and Stroop), a *z*-score

transformation based on published estimates of sample means and standard deviations was used, i.e., (x-mean)/sd. The standard scores and *z*-scores were calculated so that in the case of every variable, a higher score indicated better performance and a lower score indicated poorer performance.

The norms used for standard score and *z*-score transformations were the following: Oral Fluency (Lezak, 1983); CVLT (Delis et al., 1987); Trail Making Part B, Finger Tapping, Grooved Pegboard, Grip Strength (Bornstein, 1985); Boston Naming Test (Goodglass and Kaplan, 1983); WAIS-R Digit Symbol, WAIS-R Digit Span, WAIS-R Block Design (Wechsler, 1981); Complex Figure (Lezak, 1983); REMT (Lezak, 1983; norms for Trial 1 of the Rey Auditory Verbal Learning Test); PASAT (Roman, Edwall, Buchanan, and Patton, 1991); Wisconsin Card Sort (Heaton, 1981); and Stroop (Jensen,1965).

Next, the 19 variables were intercorrelated to look for multicollinearity. The criterion for acceptability into the final set was that no variables could be correlated at 0.80 or greater. Only two variables were intercorrelated above 0.80. These were both from the Wisconsin Card Sort Test: Perseverative Responses and Total Errors. These two scores were averaged to yield one dependent variable from the Wisconsin Card Sort Test. These 18 variables are shown in Table 5.

The final set of 18 dependent variables was used to test if the asymptomatic group performed significantly below the controls on any single measure. Multiple *t*-tests were conducted with the a priori prediction of direction allowing for a one-tailed test of

significance (see Table 5). The between-group comparisons revealed that the asymptomatic subjects were significantly ($p<.05$) lower than the homosexual controls on two tests, Oral Fluency and REMT (Number of Words Recalled on Trial 1). There was a trend ($.05<p<.10$) towards significance on four other tests: Boston Naming Test, WAIS-R Block Design, and Complex Figure Copy. We also noted an unexpected tendency for the control subjects to have lower scores on the PASAT Trial 1 than the Walter Reed 1 and 2 Group subjects.

An average *z* score for the 18 dependent variables was calculated to provide a global cognitive measure. The difference between groups on this test was also tested by a one tailed *t*-test. The groups were not different on the global *z*-score measure.

Logistic regression analysis was conducted using the two significant neuropsychological predictors (REMT and Oral Fluency) to predict the likelihood of membership in the asymptomatic group versus the homosexual control group. This analysis resulted in an estimated logistic regression model that correctly classified 40 of the 58 subjects (69%), including 22 of the 29 controls (76%) and 18 of the 29 asymptomatics (62%).

Group Comparisons on All of the Remaining Neuropsychological Measures, Information Processing Measures, and Mood Scales.

For the sake of completeness, the other 55 variables from the neuropsychological battery (scores from the WAIS-R Information

Table 5

Means and Standard Deviations of Z Scores for the 18 Neuropsychological Variables and p Values in One-Tailed T-Test Comparisons between Homosexual Controls (n=29) and WR 1 & 2 (n=29)

	Controls		WR 1 & 2		
Measure	Mean	s.d	Mean	s.d	t value
Language					
Oral Fluency total	1.00	0.99	0.51	1.23	1.69*
Boston Naming total correct	0.20	0.96	-0.60	1.34	1.31
Attention					
Digital Span	0.44	0.89	0.45	0.96	-0.05
Visual-Spatial					
Block Design	0.63	0.88	0.25	0.99	1.53
Complex Figure,Copy	0.53	1.17	0.03	1.64	1.35
Memory					
CVLT-Total 1-5	-0.67	1.03	-0.96	1.24	0.97
CVLT-Rec % const	-0.79	0.77	-0.83	1.00	0.15
REMT-Trial 1	-0.24	0.95	-0.91	1.12	2.44*
Complex Figure immediate	0.01	0.84	-0.12	1.13	0.49
Psychomotor					
Trail Making Part B-time	-0.25	0.93	-0.30	1.37	0.14
Digit Symbol	0.28	0.65	0.31	0.86	-0.18
Motor					
Tapping, ND[1]	-0.08	1.25	-0.13	0.95	0.18
Pegboard, ND[1]	0.00	0.55	0.01	0.53	-0.10
Grip Strength, ND[1]	-0.14	0.91	-0.12	1.08	-0.07
Speed of Processing					
PASAT, Trial 1	-1.73	2.46	-0.83	1.78	-1.61
Stroop, Color time	0.31	0.80	0.20	1.01	0.47

Table 5 (cont.)

Mental Flexibility					
Stroop, Color-Word time	0.06	1.11	-0.11	1.06	0.59
Wisconsin Card Sort composite*	-0.02	1.05	-0.21	1.22	0.62
Average Z Score	-0.05	0.48	-0.18	0.62	0.94

*See text for an explanation

[1] Non Dominant * $p < .05$

Test, WAIS-R Vocabulary Test, and the NART-R IQ were not re-analyzed), the 6 variables from the information processing tests (the Sternberg data were analyzed separately), and the 6 scores from the mood scales (the results from the Beck Depression Inventory and the POMS Depression Scale were not re-analyzed) were compared using a one-tailed *t*-test with $\alpha = .05$ testing the hypothesis that the asymptomatic group performed more poorly than the control group. Data from only 27 of the asymptomatic subjects were included in the analysis of hits and false alarms from the vigilance task because one subject fell asleep and the other terminated the task early.

None of the mood scales or the six information processing measures showed a significant between-group difference. Only one of the neuropsychological scores showed a significant between-group difference: the number correct on Trial 1 of List A of the CVLT (7.3 items correct for the controls versus 6.7 items for the asymptomatics). Inclusion of this additional neuropsychological measure in the logistic regression analysis, along with REMT and Oral Fluency, did not increase the likelihood of correct classification above the 69% reported for using the two predictors only.

Two-way (group by response type) ANOVAs were conducted on the slope and intercept data of the Sternberg Task. Neither analysis showed significant main effects of groups or interactions involving the group variable. Analysis of the accuracy scores is discussed below.

Analysis of Information Processing Tests Using Repeated Measures Design

In the initial examination of between-group differences described above, the average correct reaction time and the average percent correct were examined for the running difference task and the matrix task. The purpose of these analyses was to treat the information processing measures and the neuropsychological measures in the most comparable manner possible. However, these two tasks had an additional factor that is not present in the neuropsychological tests: practice (trials). The two groups could demonstrate different effects of practice on performance that would not be reflected in a significant between-group difference if the groups started at different performance levels. Thus, ignoring the effect of practice on the performance of these tasks could miss valuable information on differences in learning rates.

Because two dependent measures were obtained for each task, the correlation between the measures was obtained to determine if univariate or multivariate analyses should be used. For this analysis and all subsequent analyses involving the running difference task and the matrix task, Trial 1 was treated as a warm-up trial and not included in any analyses. The correlation between the percent correct and the average correct reaction time was calculated for each trial for all 58 subjects. Four of the 14 trials were significantly correlated for the matrix task although none of the running difference trials were significantly correlated.

A MANOVA was conducted on the data from the matrix task. Only the main effect of trial was significant [Wilks' Lambda (26, 1454)=2.71, $p<.0001$]. The effect of practice was expected and uninteresting in its own right.

A two-way (group by trial) repeated measures ANOVA was conducted on each dependent variable of the running difference task. The Huynh-Feldt adjusted degrees of freedom (Huynh and Feldt, 1976) were used in this and all subsequent univariate repeated measures analyses conducted on the running difference task. For both variables, only the main effect of trial was significant [$F(3.31, 11.41) = 639.14$, $p=.002$ for correct reaction time and $F(15.27, 3.34) = 186.96$, $p<.0001$ for the percent correct]. Again, the effect of practice was anticipated and is uninteresting in its own right.

The distributions for the correct reaction time data for both the matrix task and the running difference task were noted to be somewhat skewed and large individual differences in performance were observed. To select a method for dealing with outliers, Ratcliff (1993) was consulted. None of Ratcliff's (1993) examples appeared to fit the data well and it was difficult to select among the recommended techniques. The log transformation was eventually chosen, in part to coincide with common practice and in part to follow Ratcliff's recommendations. Thus, the MANOVA conducted on the matrix task and the ANOVA conducted on the correct reaction time data from the running difference task were calculated again after performing a log transformation of the data. In both cases

the transformation did not affect the interpretation of the data; only the main effect of trial was significant.

The percent correct for the Sternberg task can be analyzed in a manner similar to that of the matrix task and the running difference task; subjects received two blocks of five trials at each of the three positive set sizes. The last block at each level was submitted to a four-way ANOVA [group by trial by positive set size by response (*Yes* versus *No*)]. The main effects of positive set size [$F(1.49, 83.17)=9.92$, $p=.0006$] and response type [$F(1,56)=33.80$, $p<.0001$] were significant. As the positive set size increased, the percentage of correct responses decreased (97.9, 97.8, and 97.1 for set sizes 2, 3, and 4, respectively) although minimally. The subjects emitted fewer correct *Yes* responses than *No* responses (96.8% versus 98.4%). The Group by Trial interaction also was significant ($F(2.44, 136.42)=2.88$, $p=.0488$). Visual inspection of the data indicated that the control subjects showed little change over trials while the asymptomatic subjects generally improved with practice. The asymptomatic subjects did, however, start from a slightly poorer level of performance than the control subjects.

Distribution tests. The analyses described above test differences in the means of the two distributions. Although the means of the distributions do not differ, conceivably the distributions themselves may differ. For example, the skewness of one group may be greater than the other. To test this hypothesis, Komogorov-Smirnov two-sample distribution tests were conducted on

the average percentage correct and the average correct reaction time for both the running difference task and the matrix tasks. None of the four analyses showed a statistically significant difference between the two groups.

Relation Between T4 Cell Count and Test Battery Scores

Many recent studies have attempted to find a relation between laboratory markers of HIV infection and performance on assessment instruments. The means and standard deviations for the two groups on a number of laboratory markers used in the evaluation of HIV are given in Table 6. We calculated the correlation between performance on the 17 raw variables of the neuropsychological battery shown in Table 5 and the T4 count. Instead of calculating the correlation between the composite Wisconsin Card Sort measures (the 18th variable in Table 5), we calculated the correlation between each of the measures (perseverative responses and total errors) and T4 cell count. We also calculated the correlation between the T4 cell count and each of the ten measures of the information processing battery. None of the correlations for the information processing tests was significant and only two of the variables from the neuropsychological tests (REMT, number correct on Trial 1 and the Boston Naming Test, total correct) were significant (r=.3518, p<.01 and r=.2863, p<.05, respectively).

Table 6

Means for Both Groups on Five Laboratory Markers Used in the Assessment of HIV Disease

	Group			
	Control		WR 1 and 2	
Measure	mean	s.d	mean	s.d
T4 (CD4) Abs #	868.00	(296)	629.10	(222)
HGB	15.01	(0.70)	14.79	(0.80)
Albumin	4.81	(0.24)	4.52	(0.24)
RPR (Proportion Reactive)*	0/11		1/9	
B12 (pg/ml)	511.00	(281)	577.21	(310)

*The decision to include this test in the physical examination was made approximately half-way through the study. Consequently, it was only administered to 11 of the control subjects and 9 of the asymptomatic subjects.

DISCUSSION

This study had two main purposes. The first was to identify the stage at which cognitive impairments can be detected in individuals infected with the HIV virus. More specifically, the question was "At what stage of the Walter Reed Classification system are significant cognitive decrements observed compared to appropriate controls?" The second purpose of this study was to compare the relative sensitivity of information processing tests and neuropsychological tests for detecting HIV decrements. We will discuss the second purpose first.

Which Type of Test is More Sensitive?

The neuropsychological battery indicated that subjects in the Walter Reed Stages 1 and 2 Group had significantly poorer verbal episodic memory as measured by the University of Southern California Repeatable Episodic Memory Test, Trial 1 and poorer verbal fluency as measured by the Oral Fluency Test when compared to the homosexual control group. The information processing tests showed no significant between-group differences. Thus, the proportion of measures indicating a significant between-group difference was somewhat better for the neuropsychological battery (2 of 18 measures) than for the information processing battery (0 of 11 measures).

Are There Any Differences?

The primary purpose of this study was to determine when cognitive decrements can first be observed. Because the symptomatic group was lost after reclassification, this question effectively was changed to "Are there any differences between asymptomatic subjects and the controls, and, if so, what are the nature of the decrements?"

Neuropsychological tests. Subjects in the Walter Reed Stages 1 and 2 Group had significantly ($p<0.05$) poorer verbal episodic memory as measured by the University of Southern California Repeatable Episodic Memory Test, Trial 1 and verbal fluency as measured by the Oral Fluency Test compared to the homosexual control group. The Walter Reed 1 and 2 Group was slightly more than two-thirds of a standard deviation below the homosexual control group on the University of Southern California Episodic Memory Test and half a standard deviation below them on the Oral Fluency measure.

Given the number of neuropsychological tasks involved, these findings do not provide strong evidence of neurocognitive decrements in HIV+ asymptomatic subjects. If it were not for the fact that the literature shows quite consistently that tests of verbal episodic memory are reduced in patients who have ARC/AIDS, then it might be possible to dismiss these findings as either due to chance or as purely trivial. But a number of studies have reported memory decrements associated with HIV (Bornstein et al., 1992; Gibbs, Andrewes, Szmukler, Mulhall, and Bowden, 1990; Goethe

et al., 1989; Wilkie et al., 1990). The existing literature and these results indicate that memory may be one of the most sensitive neurocognitive functions to HIV infection.

Research on the sensitivity of verbal fluency to HIV infection is not as consistent as the findings with regard to memory. The Verbal Fluency Test is also one of the tests included in the Multicenter AIDS Cohort Study longitudinal research on the progression of HIV infection. Left frontal lobe structures underlie verbal fluency performance and the particular sensitivity of frontal lobe structures to perturbations with HIV infection would certainly lend some support to this as being an early affected neurocognitive function.

Neither the episodic memory nor the verbal fluency results demonstrated in this study can be completely accounted for by premorbid differences. First, we must again point out the two groups were not significantly different in terms of their NART-R, WAIS-R Vocabulary, or WAIS-R Information scores, three measures that are considered reflections of premorbid functioning. Second, the correlations between NART-R and either the episodic memory measure or the oral fluency measures were not large. For the groups combined, the correlation between University of Southern California Repeatable Episodic Memory Test 1 and NART-R was 0.25 and between oral fluency and NART-R was 0.39.

There were additional trends in the neuropsychological data that should be mentioned but not overly emphasized. The Walter Reed Stage 1 and 2 Group subjects showed a trend ($0.05<p<0.10$)

towards poorer performance on the Boston test of confrontational naming, and on two visual spatial tests, namely WAIS-R Block Design and copying of the Complex Figure. Each of these tests has been found impaired in patients with AIDS. The trend in the present data suggests that future studies of asymptomatic individuals might consider including these tests.

Information processing tests. The information processing tests showed one significant difference between the asymptomatic subjects and the control subjects on the Sternberg Task. The Group by Trials interaction on the percent correct variable reflected a greater improvement in performance by the asymptomatic group, which started from a lower accuracy level than the control group. Because there was no main effect of group, no strong conclusions should be made concerning this result. No significant differences in the distribution of the reaction times were found between the two groups on either the running difference task or the matrix task despite the promising results by Mapou et al. (1993).

Methodological Issues

As noted in the Introduction, studies examining the effects of HIV on cognitive processes have many potential methodological problems. We attempted to control or eliminate these problems by rigorous and repeated screening for substance abuse, a history of psychiatric problems, the use of psychoactive medications, and neurological problems. We only enrolled subjects whose native language was English or who acquired English by age 6.

Additionally, we matched the groups on a number of demographic variables. Specifically, the five groups in the original design were carefully matched on age, education, and ethnicity. Such matching is a time-consuming and expensive process that resulted in excluding many potential subjects and significantly narrowed the pool of potential candidates.

Despite extensive efforts to screen out major confounding variables, the homosexual control group did differ significantly on some neuropsychological tests and measures of information processing from the heterosexual control group. Thus, we had to determine which of the two groups was the appropriate control group for comparison purposes. The heterosexual control group was included to allow comparisons to data obtained from military personnel. The homosexual control group was included to control for lifestyle variables, such as alcohol and drug use, that are prevalent among sexually active homosexual men at risk for HIV infection. These variables are difficult to measure directly, but they can significantly influence cognitive performance. Thus, without controlling for the influence of these variables and by looking only at a control group of HIV- heterosexual males, the cognitive changes associated with HIV in Walter Reed Stages 1 and 2 may be overestimated. For example, if the Walter Reed 1 and 2 Group had decreased cognitive performance compared to heterosexual controls, the decrements might be attributed to HIV infection. In fact, however, such decrements might be caused by other variables that can affect cognitive function and are prevalent in the urban,

homosexual male population that is at risk for HIV infection. Thus, we selected the homosexual control group to control for possible lifestyle differences that could affect performance on the cognitive tests.

Possible Interpretation Difficulties

The selection of the homosexual control group as the appropriate reference group did not eliminate, however, all of the data interpretation problems. The sample comprising the homosexual control group and the loss of the symptomatic group caused by the reclassification may have resulted in an underestimation of the HIV-related cognitive deficits. Each of these will be discussed in turn.

The homosexual control group may have had mild-to-moderate neurocognitive decrements as indicated by their performance on a number of variables compared to the heterosexual controls, including the PASAT. The PASAT is used as a measure of speed of information processing. It was designed to test the effects of head injury and to predict the ability to return to work after head injury (Gronwall, 1977). The homosexual control group performed significantly more poorly than the heterosexual control group and the Walter Reed Stages 1 and 2 Group. The homosexual controls were 1.7 standard deviations below average, whereas the asymptomatic group was 0.8 standard deviations below average. Since the control group was confirmed to be HIV negative, the PASAT decrements must have been caused by some other variable or variables. The fact that

the homosexual control group was on the average performing at a mild-to-moderately impaired range raises the possibility that this group might have included subjects who suffered head injuries that went undetected in the screening.

We should also note that the homosexual control group reported somewhat heavier drinking than the HIV+ group (an average of 1.8 drinks per occasion for the homosexual controls and 1.3 for the asymptomatic group, a difference that was not significant for α=.05 but was significant for α=10 (p=0.06). Previous research has shown that the amount people drink per occasion is significantly related to decreased performance on certain cognitive tests (Parker, Parker, and Harford, 1991). In the present study we screened out subjects who reported drinking amounts in the range considered hazardous for cognitive functioning. However, the results of the drug toxicological screening indicated that self-reported substance abuse may have been a problem in this study; many subjects who reported no drug use actually tested positive. The main point is that the reduced scores on certain neuropsychological tests of the homosexual control group compared to the heterosexual control group raises the possibility that this group has problems other than HIV that affected their neurocognitive performance and that might not be problems to the same extent in the Walter Reed 1 and 2 Group.

The second problem concerned the loss of the Walter Reed Stages 3 and 4 Group and the Walter Reed Stage 5 Group. These groups were expec ed to exhibit significant decrements in certain areas of neurocognitive functioning compared to controls and were

to provide an endpoint of comparison. The Walter Reed 1 and 2 Group was expected to fall somewhere between the controls and the Walter Reed 3 and 4 Group. This was the type of design used by Grant and his coworkers (Grant et al., 1987) who report a pattern of performance suggestive of cognitive decrements in HIV+ asymptomatic persons who were not selected to be symptom free. In that paper there were no significant differences between the controls and the HIV+ asymptomatic group. The pattern that suggested HIV involvement was based on a significance test across four groups--controls, asymptomatic, ARC, and AIDS--and a pattern of decreasing performance in the means across the four groups on two tests, PASAT and the Category test.

Because the present study lost its more advanced stage group with reclassification, we were unable to determine internal endpoints to define the areas of cognitive decrements. The possible stepwise progression of deficits could not be examined. If there were mild decrements in the Walter Reed Stages 1 and 2 Group, but these decrements were not significant when compared to the control group, the true effects might be missed, unless there was a very large sample.

Conclusion

One of the major outcomes of this study may be a greater appreciation of two methodological problems: the choice of a control group and substance abuse assessment. The homosexual control group appears to be necessary to control for lifestyle

differences when studying HIV+ homosexual men. Nevertheless, for some comparisons an additional heterosexual control group may be appropriate. Thorough and repeated assessment of substance use must be included as screening measures. The screening should include as complete a determination as possible of the types of substances used, the frequency of use, and how long they have been used. If a toxicological screen is administered, it must be administered just before the cognitive assessment.

In conclusion, the present study provides some modest support for further research on the question of cognitive decrements in asymptomatic HIV-infected persons. The assessment battery should include neuropsychological instruments measuring verbal episodic memory, word fluency, and visual spatial functions as well as selected information processing tasks, and at least one measure of premorbid functioning.

REFERENCES

Blair, J.R., and Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist, 2*, 129-136.

Bornstein, R. (1985). Normative data on selected measures from a nonclinical sample. *Journal of Clinical Psychology, 41*, 651-659.

Bornstein, R., Nasrallah, H., Para, M., Fass, R., Whitacre, C., and Rice, R. (1991). Rate of CD4 decline and neuropsychological performance in HIV infection. *Archives of Neurology, 48*, 704-707.

Bornstein, R., Nasrallah, H., Para, M., Whitacre, C., Rosenberger, P., Fass, R., and Rice, R. (1992). Neuropsychological performance in asymptomatic HIV infection. *The Journal of Neuropsychiatry and Clinical Neurosciences, 4*, 386-394.

Delis, D.C., Kramer, J.H., Kaplan, E., and Ober, B.A. (1987). *California Verbal Learning Test research edition manual*. New York: The Psychological Corporation.

Folstein, M., Folstein, S., and McHugh, P. (1975). "Mini-mental state." *Journal of Psychiatric Research, 12*, 189-198.

Gibbs, A., Andrewes, D., Szmukler, G., Mulhall, B., and Bowden, S. (1990). Early HIV-related neuropsychological impairment: Relationship to stage of viral infection. *Journal of Clinical and Experimental Neuropsychology, 12*, 773-780.

Goethe, K., Mitchell, J., Marshall, D., Brey, R., Cahill, W., Leger, G., Hoy, L., and Boswell, R. (1989). Neuropsychological and neurological function of human immunodeficiency virus seropositive asymptomatic individuals. *Archives of Neurology, 46,* 129-133.

Goodglass, H., and Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia: Lea & Febinger.

Grant, I., Atkinson, J., Hesselink, J., Kennedy, C., Richman, D., Spector, S., and McCuthchan, J. (1987). Evidence for early central nervous system involvement in the acquired immunodeficiency Syndrome (AIDS) and other human immunodeficiency virus (HIV) infections. *Annals of Internal Medicine, 107,* 828-836.

Gronwall, D.M.A. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills, 44,* 367

Heaton, R.K. (1981). *Wisconsin Card Sorting Test manual*. Odessa, Florida: Psychological Assessment Resources, Inc.

Huynh, H., and Feldt, L. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1,* 69-82.

Jastak, S.R., and Wilkinson, G.S. (1984). *The Wide Range Achievement Test manual* (rev. ed.). Wilmington, Delaware: Jastak Associates, Inc.

Jensen, A. (1965). Scoring the Stroop Test. *Acta Psychologica, 24,* 398-408.

Karnofsky, D., and Burchenal, J., (1949). The clinical evaluation of chemotherapeutic agents in cancer. In C. MacLeod (Ed.), *Evaluation of chemotherapeutic agents*. New York: Columbia University Press.

Krikorian, R., and Wrobel, A. (1991). Cognitive impairment in HIV infection. *AIDS, 5,* 1501-1507.

Lezak, M.D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press, p.395.

Lunn, S., Skydsbjerg, M., Schulsinger, H., Parnas, J., Pedersen, J., and Mathiesen, L. (1991). A preliminary report on the neuropsychologic sequelae of human immunodeficiency virus. *Archives of General Psychiatry, 48,* 139-142.

Mapou, R., Kay, G., Rundell, J., and Temoshok, L. (1993). Measuring performance decrements in aviation personnel infected with the human immunodeficiency virus. *Aviation, Space, & Environmental Medicine, 64,* 158-164.

Martin, A., Heyes, M., Salazar, A., Kampen, D., Williams, J., Law, W., Coats, M., and Markey, S. (1992). Progressive slowing of reaction time and increasing cerebrospinal fluid concentrations of quinolinic acid in HIV-infected individuals. *The Journal of Neuropsychiatry and Clinical Neurosciences, 4,* 270-279.

McAllister, R., Herns, M., Harrison, M., Newman, S., Connolly, S., Fowler, C., Fell, M., Durrance, P., Manji, H., Kendall, B.,

Valentine, A., Weiler, I., and Adler, M. (1992). Neurological and neuropsychological performance in HIV seropositive men without symptoms. *Journal of Neurology, Neurosurgery, and Psychiatry, 55*, 143-148.

McArthur, J., Cohen, B., Selnes, O., Kuman, A., Cooper, K., McArthur, J., Soucy, G., Cornblath, D., Chmiel, J., Wang, M., Starkey, D., Ginzburg, H., Ostrow, D., Johnson, R., Phair, J., and Polk, B. (1989). Low prevalence of neurological and neuropsychological abnormalities in otherwise healthy HIV-1-infected individuals: Results from the Multicenter AIDS Cohort Study. *Annals of Neurology, 26*, 601-611.

McNair, D., Lorr, M., and Droppleman, L. (1971). *Manual for the profile of mood states*. San Diego, Ca: Educational and Industrial Testing Service.

Nelson, H. (1982). *National Adult Reading Test (NART): Test manual*. Windsor, UK: NFER-Nelson Publishing.

Nelson, H., and O'Connell, A. (1978). Dementia: The estimation of premorbid intelligence levels using the New Adult Reading Test. *Cortex, 14*, 234-244.

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologica, 9*, 97-113.

Parker, E., Bridge, T., Ingraham, L., Eaton, E., and Heseltine, P. (1989). *Neurotoxicity battery for Phase 1 HIV drug assessment: Initial applications*. Abstract NP12, Neurological and Neuropsychological Complications of HIV Infection, Quebec City, Quebec, Canada, May 31-June 3, 1989.

Parker, E., Parker, D., and Harford, T. (1991). Specifying the relationship between alcohol use and cognitive loss: The effects of frequency of consumption and psychological distress. *Journal of Studies of Alcohol, 52*, 366-373.

Poutiainen, E., Elovaara, I., Raininko, R., Hokkanen, L., Valle, S.-L., Lahdevirta, J., and Iivanainen, M. (1993). Cognitive performance in HIV-1 infection: Relationship to severity of disease and brain atrophy. *Acta Neurologica Scandinavica, 87*, 88-94.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*, 510-532.

Rogers, J., Howard, K., and Vessey, J. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin, 113*, 553-565.

Roman, D., Edwall, G., Buchanan, R., and Patton, J. (1991). Extended norms for the Paced Auditory Serial Addition Task. *The Clinical Neuropsychologist, 5*, 33-40.

Rosci, M., Pigorini, F., Bernbei, A., Pau, F., Volpini, V., Merigliano, D., and Meligrana, M. (1992). Methods for detecting early signs of AIDS dementia complex in asymptomatic HIV-1-infected subjects, *AIDS, 6*, 1309-1316.

Saykin, A., Janssen, R., Sprehn, G., Kaplan, J., Spira, T., and Weller, P. (1988). Neuropsychological dysfunction in HIV-infection: Characterization in a lymphadenopathy cohort. *International Journal of Clinical Neuropsychology, X*, 81-95.

Sinforiani, E., Mauri, M., Bono, G., Muratori, S., Allessi, E.,

and Minoli, L. (1991). Cognitive abnormalities and disease progression in a selected population of asymptomatic HIV-positive subjects. *AIDS, 5,* 1117-1120.

Squire, L., and Zouzounis, J. (1988). Self-ratings of memory dysfunction: Different findings in depression and amnesia. *Journal of Clinical and Experimental Neuropsychology, 10,* 727-738.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276-315.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale--Revised (WAIS-R).* New York: The Psychological Corporation.

Wilkie, F., Eisdorfer, C., Morgan, R., Lowenstein, D., and Szapocznik, J., (1990). Cognition in early human immunodeficiency virus infection. *Archives of Neurology, 47,* 433-440.

Wilkie, F., Morgan, R., Fletcher, M., Blaney, N., Baum, M., Komaroff, E., Szapocznik, J., and Eisdorfer, C. (1992). Cognition and immune function in HIV-1 infection. *AIDS, 6,* 977-981.